

Lossless Intra Coding in HEVC with Adaptive 3-tap Filters

Saeed Ranjbar Alvar and Fatih Kamisli

Department of Electrical and Electronics Engineering, Middle East Technical University, Turkey

Emails: Saeed.alvar@metu.edu.tr and Kamisli@eee.metu.edu.tr

Abstract—In pixel-by-pixel spatial prediction methods for lossless intra coding, the prediction is obtained by a weighted sum of neighbouring pixels. The proposed prediction approach in this paper uses a weighted sum of three neighbor pixels according to a two-dimensional correlation model. The weights are obtained after a three step optimization procedure. The first two stages are offline procedures where the computed prediction weights are obtained offline from training sequences. The third stage is an online optimization procedure where the offline obtained prediction weights are further fine-tuned and adapted to each encoded block during encoding using a rate-distortion optimized method and the modification in this third stage is transmitted to the decoder as side information. The results of the simulations show average bit rate reductions of 12.02% and 3.28% over the default lossless intra coding in HEVC and the well-known Sample-based Angular Prediction (SAP) method, respectively.

I. INTRODUCTION

Video coding standards such as the state of art High Efficiency Video Coding (HEVC) [1] and widely used H.264/AVC [2] support both lossy and lossless compression. In both lossy and lossless compression modes, prediction is performed in a block based approach and then the difference between the original block and the predicted block (residual block) is further processed depending on the mode of compression and the input configurations.

In lossless coding, transform and quantization are skipped and the prediction residual block is directly entropy coded. In the mentioned standards, since the residual is obtained from a block based prediction which cannot provide a sufficiently well prediction for pixels away from the prediction boundaries, the energy of the residual block is high. In order to decrease the energy of the residual block, two set of approaches are proposed. In the first set [3], [4], [5], [6], [7], the residual block is post processed such that its energy is lowered. In the second set of approaches [8], [9], [10], [11], [12], the prediction is obtained by using pixel-by-pixel spatial prediction instead of block-based prediction. The details of the spatial prediction methods based on pixel-by-pixel predictions will be discussed in the following section.

In this paper, a pixel-by-pixel spatial prediction method which uses three neighbouring pixels, similar to the algorithm discussed in [12], is proposed. However, while [12] obtains prediction weights offline from training sequences, this paper uses an online method to adapt the prediction weights to the

content of each encoded block during encoding, which can provide further coding gains.

This paper is organised as follows. In section II the spatial prediction methods based on pixel-by-pixel prediction will be discussed. Section III discusses the details of intra lossless coding with 3-tap filters. In section IV the proposed algorithm is introduced. Next section describes the details of the implementations the analyzes the performance of the implemented algorithms. Finally, the paper is concluded in section VI.

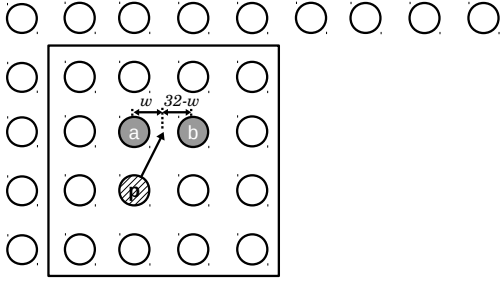
II. SPATIAL PREDICTION METHODS BASED ON PIXEL-BY-PIXEL PREDICTION

When the transform step is skipped in lossless intra coding, the block-based spatial prediction becomes less effective since some block pixels are predicted from distant reference samples and there is no transform step that can compensate for this inefficient prediction. However, since the transform is skipped in lossless coding, a pixel-by-pixel spatial prediction approach can now be used instead of a block-based approach for more efficient prediction.[12] Sample-based Angular Prediction (SAP) [8] is an example of a pixel-by pixel prediction. In SAP, Modes Planar and DC are kept as HEVC's default modes and the angular modes are modified. In the directional modes the prediction angle and the formula utilized to form the output is the HEVC's angle and formula, however, there is one key difference between these two. HEVC projects the current pixel into the reference samples (i.e. neighbor pixels of the block), on the other hand, SAP uses the same direction but the current pixel is projected to the immediate neighboring pixels. After the projection, the two closest pixels to the location of the projection are linearly interpolated and the value of the predicted pixel is obtained using the following formula.

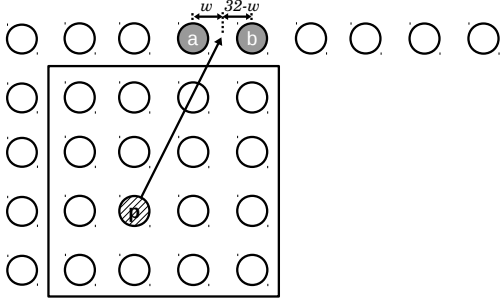
$$p = ((32 - w) \cdot a + w \cdot b + 16) >> 5. \quad (1)$$

Here ">>" indicates a bit shift operator, a and b represent the reference samples (closest pixels to the projected location), and $32 - w$ and w represent 5-bit integer interpolation weights, which are determined by the angle or prediction mode [13]. Figure 1 is an example of the projection of a pixel in SAP and HEVC.

Similar algorithm to SAP is adaptive directional SAP (AD-SAP)[9]. In this approach the prediction is similar SAP[8] but encoder may change the direction of the prediction if it can remove the spatial redundancy more efficient than the ordinary SAP.



(a) Sample-based Angular Prediction (SAP) proposed for HEVC in [8]. Block pixel to be predicted is projected on the immediately above row (or column, depending on mode) pixels (reference samples) along an angular direction. Prediction p is obtained from linear interpolation of two closest reference samples a and b .



(b) Block-based angular prediction in HEVC. Block pixel to be predicted is projected on the block neighbor pixels (reference samples) along an angular direction. Prediction p is obtained from linear interpolation of two closest reference samples a and b .

Fig. 1: difference of the projection in SAP (a) and in HEVC (b)

One other example for algorithms of this type is Piecewise DC prediction where the average of the left and above pixels adjacent to the current pixel in the prediction unit (PU) is utilized for obtaining the prediction [10].

Another approach based on pixel by pixel prediction method is discussed in [12]. In this method three neighboring pixels are used for prediction according to a two-dimensional correlation model. The details of this algorithm is discussed in the following section.

III. LOSSLESS INTRA CODING WITH 3-TAP FILTERS

A. 3-tap filtering approach

In the most of the above mentioned algorithms, two neighboring pixels are used to obtain the prediction. The algorithm discussed in [12] adds the third pixel and the weighted sum of these three pixels is the final value of the prediction for all intra modes. In order to represent the directionality of intra modes in a more efficient way, these three neighbors are not fixed and they change depending on the mode of the prediction. In this method, the value of the prediction is obtained using the following equation:

$$p = \rho_{1,k} \cdot a + \rho_{2,k} \cdot b + \rho_{3,k} \cdot c. \quad (2)$$

a , b and c are the locations of neighboring pixels to be used in the prediction, see (Fig 2) and $\rho_{1,k}$, $\rho_{2,k}$ and $\rho_{3,k}$ are weights corresponding to mode k .

B. Prediction weights

The challenging issue with the prediction based on 3-tap filters is the value of the weights. The weights are obtained from a training sequence¹ using a two-stage optimization process. In the first stage, an iterative minimum squared-error(MSE) method is used, which finds the weights that minimize the mean squared prediction error over the training sequence. Let's B_{opt} be the resulting bitrate from the parameters obtained from the first stage and $\rho_{1,k}$, $\rho_{2,k}$ and $\rho_{3,k}$ be the prediction weights of intra mode k . Using B_{opt} and $\rho_{1,k}$, $\rho_{2,k}$ and $\rho_{3,k}$, the second stage of optimization is performed as following [12]:

- 1) Let $k = 0$ and $B_{best} = B_{opt}$.
- 2) Generate 6 candidates for prediction weights of mode k ($\rho_{1,k,i}$, $\rho_{2,k,i}$ and $\rho_{3,k,i}$), run HEVC coder by replacing mode k 's weights with the candidates and record the resulting bitrates B_i . Find candidate with

TABLE I: Candidate prediction weights

(Candidate) i	$\rho_{1,k,i}$	$\rho_{2,k,i}$	$\rho_{3,k,i}$	Bitrate B_i
1	$\rho_{1,k} + 1$	$\rho_{2,k} - 1$	$\rho_{3,k}$	B_1
2	$\rho_{1,k} - 1$	$\rho_{2,k} + 1$	$\rho_{3,k}$	B_2
3	$\rho_{1,k} + 1$	$\rho_{2,k}$	$\rho_{3,k} - 1$	B_3
4	$\rho_{1,k} - 1$	$\rho_{2,k}$	$\rho_{3,k} + 1$	B_4
5	$\rho_{1,k}$	$\rho_{2,k} + 1$	$\rho_{3,k} - 1$	B_5
6	$\rho_{1,k}$	$\rho_{2,k} - 1$	$\rho_{3,k} + 1$	B_6

smallest bitrate, $i^* = \arg\min_i B_i$. If $B_{i^*} < B_{opt}$, update bitrate $B_{opt} = B_{i^*}$ and prediction weights for mode k , $\rho_{1,k} = \rho_{1,k,i^*}$, $\rho_{2,k} = \rho_{2,k,i^*}$ and $\rho_{3,k} = \rho_{3,k,i^*}$.

- 3) If $k < \text{number of modes}$, increment k by one and go to step 2. If $k = \text{number of modes}$, check if this iteration over all intra modes improved bitrate, i.e. $B_{opt} < B_{best}$. If so, go to step 1, otherwise finish.

While MSE and bitrate are in general coherent, reduction of MSE does not necessarily results in lower bitrate, therefore a second-stage optimization is performed to further fine-tune the weights to achieve minimum bitrate.

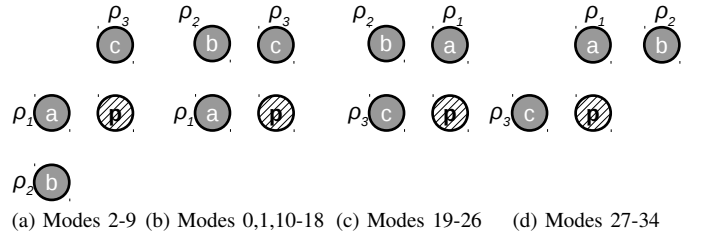


Fig. 2: Neighbor pixels used for prediction in the 3-tap filtering method according to intra modes of HEVC. Intra modes 2-9 use neighbor pixels shown in (a), planar, DC modes (0,1) as well as intra modes 10-18 use neighbor pixels shown in (b), intra modes 19-26 use neighbor pixels shown in (c) and intra modes 27-34 use neighbor pixels shown in (d).

¹A training sequence was formed from several images in the JPEG-XR image test set [14].

IV. LOSSLESS INTRA CODING WITH ADAPTIVE 3-TAP FILTERS

According to the results discussed in [12] the lossless intra coding using 3-tap filters provides substantial bitrate savings compared to HEVC lossless coding. In addition, it is shown that the second stage of the optimization discussed in III-B provides additional average 0.7% bit rate reduction over the gains obtained from the first step. This proves that by further adjusting the weights of the filters we have the chance to obtain more accurate prediction. In other words, if we can modify the offline parameters of the 3-tap filters adaptive to the content of the current PU, we can reach a better prediction for that PU.

As discussed earlier, the offline weights in [12] are obtained from the data of a training set and those weights stay constant during the encoding and decoding procedure. The algorithm that is proposed in this paper keeps the prediction as in [12] but tries to use a technique similar to the second stage of the optimization discussed in III-B by modifying the ρ parameters during the Rate Distortion Optimization (RDO) process. These adaptive weights help in finding the more accurate weights for the prediction.

HEVC reference software, HM [15], uses a three-step RDO method to find the optimal intra mode. In the first step, it returns N most promising modes out of 35 modes using Hadamard cost for RDO search. The N depends upon the PU size. The value of N can be {8, 8, 3, 3, 3} for PU size = {4x4, 8x8, 16x16, 32x32, 64x64}, respectively. Most Probable Modes (MPM) are then added to promising modes. In the second-step, RDO mode selection is performed and it returns the best mode based on rate distortion (RD) cost. Finally, the third step decides the best Transform Units (TU) partitioning for the current PU given the mode selected in the second-step [16]. The first step is performed in order to avoid RDO mode selection to be tested for every possible intra mode. Hadamard Transform is chosen to provide more proper candidates for RDO mode selection process by simulating what transforms of HEVC do during RDO but with a much simpler approach and with less number of operations. Since there is no transform in lossless case, using Hadamard cost is not efficient here. As a result, the first point that is suggested in this paper is to change the Hadamard cost to Sum of Absolute Difference between the original block and the prediction block (SAD cost) and find the candidates for RDO mode selection based on SAD cost.

The main goal which is further optimizing the ρ parameters can be achieved in the first step of the described RDO process. The modified RDO process is as following: in the first step, similar to HEVC, N (stays as it is in HEVC) most promising modes out of 35 modes for a PU are obtained using the parameters in [12]. Then before entering RDO mode selection step, for each of the N candidates, 8 set of ρ parameters (1 as given in [12], 6 as discussed in III-B, and 1 randomly chosen) are assigned and then sent to a new loop similar to first stage's loop which iterates 8 times. This 8 iterations,

return 8 different SAD costs corresponding to each set of parameters. Then the best N among 8*N (N modes with 8 different set of parameters) cost that give the lowest SAD cost are proceeded to the next steps and finally the best mode and the corresponding set of parameters are resolved. It should be noted that we must signal which of those 8 sets achieved the lowest RD cost to be able to use it in other stages of the encoding and during the reconstruction in the decoder as well. Signaling the best candidate is done by a setting a 3 bit flag for each PU, which is written into the bit stream and is considered in cost calculation during the whole encoding process. Basically there are 7 modes for the flag to take (1 as given in [12] and 6 as discussed in III-B). However, since a 3 bit flag is used, the eighth candidate is suggested randomly to use the capacity of 3 bits in signaling the candidates as much as possible. Since 3 bits of redundant data is a considerable ratio of the data size in 4x4 and 8x8 blocks, the improvement of the modified parameters cannot compensate the cost of three bits in most of the cases. As a result, the proposed algorithm is applied for the remaining block sizes and 4x4 and 8x8 PUs are predicted using the offline parameters given in [12]. Algorithm 1 summarizes the proposed algorithm.

Algorithm 1 Proposed algorithm

- 1: Start the RDO procedure
 - 2: Obtain the N most promising modes out of 35 using SAD cost
 - 3: **for** each of N candidates **do**
 - 4: obtain additional 7 set of parameters and obtain the SAD cost for each.
 - 5: **end for**
 - 6: Choose N candidates with the lowest SAD cost among 8*N candidates and update N most promising modes by setting the modes and the flags indicating which of 7 set of additional parameters exists in the list.
 - 7: Perform RDO mode selection and TU splitting stages and finalize the intra mode of PU.
 - 8: Write the information of the best mode and the corresponding flag into the bit stream.
-

V. EXPERIMENTAL RESULTS

The proposed prediction method based on adaptive 3-tap filters, method based on 3-tap filters using offline weights [12] and SAP [8] are implemented in the HM12.0 [15]. Initial 50 frames of the sequences in Classes A to F are tested in AI-Main configuration and QP=0, with the common test conditions [17]. It should be noted that training sequences used in III-B don't include any of tested sequences. Table II shows the average percentage bitrate reduction of three approaches compared to the lossless intra coding in HEVC.

From the comparison of the results, It can be seen that the proposed algorithm achieves the highest gains among the implemented algorithms. In addition, the comparison of the results of 3-tap filters with SAP shows how effective the third pixel is in removing the spatial redundancy. The results also

TABLE II: Average percentage bitrate reduction of multiple methods and the adaptive 3-tap filtering method over HEVC lossless intra coding.

	SAP[8]	3-tap filters (offline weights) [12]	adaptive 3-tap filters
Class A	8.94	15.60	15.92
Class B	5.78	8.59	8.95
Class C	6.95	8.67	9.11
Class D	8.91	11.06	11.40
Class E	10.56	14.35	14.76
Class F	12.51	12.48	13.44
Average	8.74	11.55	12.02

reveal that the proposed algorithm has improved the gains of the 3-tap filters with the offline parameters for all the classes, especially for Class F where in average 0.96% bit rate reduction is observed.

VI. CONCLUSION

In this paper a novel pixel-by-pixel spatial prediction method based on the 3-tap filters is proposed for lossless intra coding of HEVC. In the proposed algorithm, despite the conventional prediction based on 3-tap filters, the weights of the 3-tap filters are not constant and the modified value of the weights are explored adaptively during the RDO process.

The comparison of the performance of the proposed method with the HEVC's lossless gains in HM software, shows the average 12.02% bit rate reduction. In addition, the proposed algorithm improves the performance of the intra prediction based on 3-tap filters with fixed weights, up to 0.96% for some of the tested classes.

ACKNOWLEDGMENT

This research was supported by Grant 113E516 of Tubitak.

REFERENCES

- [1] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video coding (HEVC) standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.
- [2] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, July 2003.
- [3] Y.-L. Lee, K.-H. Han, and G. Sullivan, "Improved lossless intra coding for H.264/MPEG-4 AVC," *Image Processing, IEEE Transactions on*, vol. 15, no. 9, pp. 2610–2615, Sept 2006.
- [4] S.-W. Hong, J. H. Kwak, and Y.-L. Lee, "Cross residual transform for lossless intra-coding for HEVC," *Signal Processing: Image Communication*, vol. 28, no. 10, pp. 1335 – 1341, 2013.
- [5] J.-H. Kwak and Y.-L. Lee, "Secondary residual transform for lossless intra coding in HEVC," *Journal of Broadcast Engineering*, vol. 17, no. 5, pp. 734–741, 2012.
- [6] G. Jeon, K. Kim, and J. Jeong, "Improved residual DPCM for HEVC lossless coding," in *Graphics, Patterns and Images (SIBGRAPI), 2014 27th SIBGRAPI Conference on*, Aug 2014, pp. 95–102.
- [7] X. Cai and J. S. Lim, "Adaptive residual DPCM for lossless intra coding," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2015, pp. 94 100A–94 100A.
- [8] M. Zhou, W. Gao, M. Jiang, and H. Yu, "HEVC lossless coding and improvements," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1839–1843, Dec 2012.
- [9] X.-P. Xia, E.-H. Liu, and J.-J. Qin, "Improved sap based on adaptive directional prediction for hevc lossless intra prediction," *Journal of Visual Communication and Image Representation*, vol. 33, pp. 78–84, 2015.
- [10] K. Kim, G. Jeon, and J. Jeong, "Piecewise DC prediction in HEVC," *Signal Processing: Image Communication*, vol. 29, no. 9, pp. 945 – 950, 2014.
- [11] E. Wige, G. Yammine, P. Amon, A. Hutter, and A. Kaup, "Pixel-based averaging predictor for HEVC lossless coding," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, Sept 2013, pp. 1806–1810.
- [12] S. R. Alvar and F. Kamisli, "Lossless intra coding in hevc with 3-tap filters," *arXiv preprint arXiv:1601.04473*, 2016.
- [13] J. Lainema, F. Bossen, W.-J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1792–1801, Dec 2012.
- [14] "JPEG core experiment for the evaluation of JPEG XR image coding." [Online]. Available: <http://documents.epfl.ch/groups/g/gr/gr-eb-unit/www/IQA/Original.zip>
- [15] "HM 12.0 reference software." [Online]. Available: <https://hevc.hhi.fraunhofer.de/trac/hevc/browser/tags/HM-12.0>
- [16] Z. Sheng, D. Zhou, H. Sun, and S. Goto, "Low-complexity Rate-Distortion Optimization algorithms for HEVC intra prediction," in *Multimedia Modeling*, ser. Lecture Notes in Computer Science, C. Gurrin, F. Hopfgartner, W. Hurst, H. Johansen, H. Lee, and N. OConnor, Eds. Springer International Publishing, 2014, vol. 8325, pp. 541–552.
- [17] F. Bossen, "Common test conditions and software reference configurations," *Joint Collaborative Team on Video Coding (JCT-VC), JCTVC-F900*, 2011.